International Environmental Modelling and Software Society (iEMSs)
2012 International Congress on Environmental Modelling and Software
Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany
R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)
http://www.iemss.org/society/index.php/iemss-2012-proceedings

Conceptual and Practical Aspects of Quantifying Uncertainty in Environmental Modelling and Decision Support

Peter Reichert

Eawag, Swiss Federal Institute of Aquatic Science and Technology
8600 Dübendorf, Switzerland
reichert@eawag.ch

Abstract: Environmental decision support intends to use the best available scientific knowledge to predict the consequences of management alternatives. This raises 3 questions:

(i) How to formally represent and quantify scientific knowledge?
(ii) How to find adequate model structures and parameter values for predicting the behaviour of environmental systems under different driving conditions?
(iii) How to implement efficient numerical procedures to actually calculate such predictions?

Approaches to address all three of these questions will briefly be discussed. With respect to (i) an intersubjective interpretation of probabilities with an extension to imprecise probabilities is suggested as the most adequate representation of scientific knowledge. Conceptual arguments in favour of this approach are discussed as well as problems of alternative approaches. To address (ii) the importance of considering input errors, model structure deficiencies, and internal stochasticity of the modelled system is emphasized, as well as handling remaining systematic errors or bias in model output adequately. This typically requires Bayesian inference and the propagation of probability distributions through deterministic or stochastic models. Under (iii) numerical difficulties of Bayesian inference are briefly mentioned as well as approaches to overcome these. Of particular importance are adaptive Markov Chain Monte Carlo methods, Approximate Bayesian Computation (ABC) techniques that make it possible to simulate from the likelihood function instead of evaluating it, and recent approaches of statistically emulating computationally demanding dynamic models. Finally, an overview of research needs is provided.

Keywords: *Representing Knowledge; Uncertainty; Intersubjective Probabilities; Imprecise Probabilities; Bias; Environmental Modelling.*

1 INTRODUCTION

Decision making is always based on an assessment of the desirability of the outcomes of decision alternatives. This requires the quantification of the preferences of the decision maker(s) and the prediction of the consequences of all decision alternatives. In environmental management estimating the consequences of decision alternatives is particularly difficult. The high complexity of natural, technical and socio-economic systems and our incomplete knowledge of their structure and function lead to a high uncertainty of such estimates. There are three reasons, why the quantification of this uncertainty is extremely important:

- In environmental management, the prediction of the consequences of decision alternatives is usually done by scientists whereas the assessment of the desirability of the outcomes has to reflect societal preferences. Key principles of scientific integrity and ethical responsibility of scientists towards the society require the scientists to communicate their uncertainty as an essential part of the scientific input into societal decision making. There are many examples in the past in which this uncertainty was not communicated and wrong predictions led to an erosion of the trust of stakeholders to predictions by scientists.

- Different decision alternatives may have similar environmental consequences. It may be relevant for the choice between such alternatives to know if differences in environmental consequences are significant or if they are smaller than the uncertainty of the predictions. If they are much smaller, the consequences may be felt to be equivalent and other criteria may gain importance for decision making. However, we should also be aware of the fact that predictions may be uncertain but, due to the dependence structure of alternatives, the difference in outcomes can still be quite certain [Reichert and Borsuk, 2005].
- Facing predictions with different degrees of uncertainty, risk attitudes may become an important element of societal decision making. In particular, risk aversion (e.g. the wish to avoid unwanted outcomes even if they have a very low probability) can be an essential criterion in societal decision making. Extreme examples of this are decisions about implementing or not implementing technologies that may have (a low probability for) strongly unwanted outcomes. Despite all scientific uncertainty, we can avoid these unwanted outcomes with certainty if we do not implement the technology (but other uncertainties may be involved with other alternatives).

Over the past decades, the awareness of the scientific community about the need of quantifying scientific uncertainty has increased considerably and many approaches to describe and quantify the uncertainty of scientific predictions have been published. However, despite fundamental differences in the underlying concepts of these approaches, there seems to be insufficient discussion about their appropriateness. Different schools developed in parallel and they follow different approaches without clear justification of the superiority of one approach over the other. It is the intent of this paper to stimulate the discussion about the appropriateness of different approaches to describe scientific knowledge and uncertainty.

Due to space limitations in this conference paper, this paper concentrates on a brief discussion of the key elements. We plan to publish an extended argumentation that will also cover the aspects of quantifying preferences to complete the conceptual basis of environmental decision support [Reichert et al., in preparation].

2 REPRESENTING AND QUANTIFYING SCIENTIFIC KNOWLEDGE

In this section the motivation for using the mathematical concept of probabilities to describe uncertain scientific knowledge is briefly reviewed (section 2.1), it is extended to imprecise probabilities (section 2.2) and compared with alternative theories that have been suggested in the literature (section 2.3).

2.1 Intersubjective Probabilities

Relative frequencies of repeated observations of a system with a random component are conceptually (this means in the limit of very large numbers of observations) described by probabilities. This is called the frequentist interpretation of probabilities. As it is evident that relative frequencies and their limits fulfill the axioms of probability, there is no debate about the adequateness of this description. For environmental systems, potential causes of such (approximately) random behaviour are influence factors that were not considered (or even not observed) or the aggregated description of the system. However, it is evident that there are more sources of uncertainty of the behaviour of a system than its intrinsic randomness. The incomplete knowledge about the inputs to the system and its structure and function induces additional uncertainty. Usually, uncertainty will be caused by a combination of uncertainty due to non-deterministic behaviour (aleatory uncertainty) and due to lack of knowledge (epistemic uncertainty). It is not a priori evident by which mathematical formalism this uncertainty or partial knowledge should be described. If we assume, however, that the degree of belief or knowledge about a possible outcome can be quantified by a single number, there are the following arguments, why also subjective belief of individual experts or intersubjective knowledge of a group of scientists should be described by probabilities:

1. **Avoiding sure loss.** We need an operational definition of "belief" or "knowledge" as this is a much more vague notion than is a frequentist probability. A straightforward way of operationalizing a belief about a statement is by indifference between lotteries about the statement. If, for a certain value of p, an individual is indifferent between the lotteries with a gain proportional to (1-p) if a statement is true and a loss proportional to p if it is false and the lottery with the negatives of these payoffs, then her or his belief in the statement to be true can be quantified by the value of p. It can be shown that if an individual agrees to operationalize her or his beliefs with such lotteries and if the individual wants to avoid sure loss (to an opponent who can choose between lotteries about which the person is indifferent), then these quantities p must fulfil the axioms of probability calculus (Ramsey – De Finetty theorem; see Howson and Urbach 1989 for a careful discussion and proof). This argumentation for individuals can easily be extended to groups what results in intersubjective rather than subjective probabilities [Gillies 1991, 2000].
2. **Formulation of conditional beliefs.** Cox [1946] (see also discussion in Snow [1998] and Colyvan [2004]) shows that conditional beliefs must follow the laws of probability if they fulfil the following requirements (Cox's theorem): (i) the degree of belief can be expressed as a real number; (ii) the degree of belief in a statement A determines the degree of belief of its negation; (iii) the degree of belief in two statements A and B depends only on the degree of belief in B and in A given B (and some implicit assumptions, see e.g. Colyvan [2004]). This argument is relevant because of the high importance of the formulation of conditional knowledge for environmental decision support.
3. **Consistency with frequentist probabilities.** As mentioned above, important contributions to uncertainty in predictions are non-deterministic behaviour of the investigated system and lack of knowledge about its mechanisms. Obviously, it depends on the system, which of these components dominates the overall uncertainty or if they are of similar importance. Non-deterministic behaviour is often assumed to be random and can thus best be described by (frequentist) probabilities. When the random outcome is realized but not yet observed, this aleatory uncertainty becomes epistemic uncertainty (lack of knowledge in this case about an outcome that has become certain). In this situation we obviously run into inconsistencies if we use a different mathematical formulation for describing aleatory than epistemic uncertainty.

It is interesting that there is a diversity of different arguments in favour of using the mathematical framework of probabilities to describe subjective beliefs as well as intersubjective knowledge (in our case of the scientific community). Nevertheless, the assumptions underlying this conclusion can be questioned (see also section 2.2) and other ways of describing epistemic uncertainty are possible (Colyvan 2008). However, it is important to question and discuss the conceptual foundation of any technique used to describe scientific uncertainty and to compare this foundation with that of probability calculus. The use of ad-hoc procedures without a conceptual foundation should be avoided.

2.2 Consideration of Imprecision

In section 2.1 it was argued that a probabilistic description seems to be ideal for describing intersubjective, scientific knowledge. However, when trying to acquire such knowledge, it becomes obvious that its unique specification is difficult. Scientists may be uncertain about their own beliefs and different scientists may quantify their beliefs differently. This makes the description of the degree of belief by a unique real number questionable. Note that this invalidates all three arguments in favour of the probabilistic approach in section 2.1. However, these arguments remain valid for the limiting case of precisely specified beliefs. For this reason, it seems meaningful to look for an extension of probability theory that reduces to probability theory for the case of precise formulations of beliefs rather than giving up the probabilistic concept completely. This is the case when replacing unique probability distributions by sets of distributions, so called imprecise probabilities (e.g. Walley 1991, http://www.sipta.org). For the special case of the so-called density ratio class of imprecise probabilities (DeRobertis and Hartigan, 1987), there were recent efforts made to improve its accessibility [Rinderknecht et al., 2011; Rinderknecht et al. 2012].

2.3 Alternative Theories

The claim of probability theory to be the unique correct framework for describing uncertainty has been challenged [Oberkampf et al., 2004; Helton and Oberkampf, 2004; Colyvan, 2008] and alternative approaches have been suggested (Helton et al 2004). The most prominent approaches are possibility theory [Zadeh, 1978; Dubois and Prade, 1988; Dubois 2006], evidence theory [Dempster, 1967; Shafer, 1976], and interval analysis [Moore, 1979]. Besides these theories, also more “informal” approaches have been developed, e.g. based on arbitrary modifications of the likelihood function [GLUE, e.g. Beven and Freer, 2001] or on the construction of uncertainty intervals based on the coverage frequency of observations [SUFI; Abbaspour et al. 2004].

The main reasons for the development of possibility theory, evidence theory and interval analysis were similar to the motivation for developing imprecise probabilities discussed in section 2.2 [Colyvan, 2008], whereas the reason for the development of GLUE and SUFI was primarily the development of a simpler approach that leads to reasonable uncertainty estimates. None of these alternative theories has a similarly good axiomatic foundation for being an ideal representation of uncertainty as probability theory. However, the development of the foundations of alternative theories should be followed and the evaluation of the optimal technique continued.

3 DESIGN OF ENVIRONMENTAL MODELS

Environmental models are designed to summarize our knowledge by representing the structure and mechanisms of real systems and to predict future behaviour typically under alternative driving conditions corresponding to different future scenarios or decision alternatives. This requires a very careful design of the models and, in particular, a careful handling of uncertainty.

3.1 Model Structure

If the model is designed to predict effects of previously unobserved changes in external driving conditions, it has to be based on an at least partially mechanistic description of the structure and function of the system described by the model. As environmental systems are extremely complex, the model must consist of a simplified description of reality. Finding the adequate level of simplification is often difficult as a compromise must be found between a model that behaves incorrectly because it lacks a description of important mechanisms and a model that has a very high prediction uncertainty because its complexity involves a large set of poorly identifiable parameters to be inferred. If possible, an approach using different models based on different concepts and levels of complexity would be the best way to address this structural uncertainty of the modelling process.

An important element that is often overlooked in traditional mass-balance based environmental models is the adequate description of non-deterministic behaviour mainly due to aggregation errors (e.g. approximation of random processes at the micro-scale by continuous, deterministic processes, spatial aggregation, aggregation of species into functional groups, etc.), influence factors that are not considered, and model structure deficiencies (e.g. neglected adaptation processes, etc.). In the absence of mechanistic knowledge about these processes, it can be important to consider the non-deterministic behaviour of the system due to such causes by internal stochasticity of the model. It is conceptually satisfying to do this by still keeping the mass balances exactly correct. This can e.g. be done by approximating random interactions between a finite number of particles by a continuous stochastic process [Wilkinson, 2011] or by making model parameters stochastic processes [Reichert and Mieleitner, 2009; Lin and Beck, 2012] rather than making deterministic mass-balance equations stochastic [e.g. Vrugt et al. 2005]. Having added internal stochasticity to a model to account for the non-deterministic behaviour of the system, it remains important to clearly separate this stochasticity from uncertainty due to lack of knowledge [Oberkampf et al., 2004], because the latter is reducible by acquiring more information about the system.

Model parameterization should be done carefully to avoid making posterior distributions unnecessarily complex [Kavetski and Kuczera, 2007] and care should be taken to avoid artefacts due to poor numerical procedures [Clark and Kavetski, 2010; Kavetski and Clark, 2010].

3.2 Model Calibration and Dealing with Model Deficiencies

In a Bayesian context, updating prior distributions of model parameters consists of conditioning the prior distribution with the observed data. Prior distributions describe knowledge about the parameter value that is not based on the (newly) observed data used for updating. As only frequentist probabilities can be validated empirically, it may often not be possible to validate the statistical assumptions underlying a model. There is a very important exception. In many cases, models are described by likelihood functions that have a frequentist interpretation. This means that given the correct parameter values, we assume that the system observations are a sample of the distribution formulated by the likelihood function. If we formulate our knowledge about the parameter values by Bayesian probability distributions, we can easily use the same, frequentist likelihood function for inference, as it seems reasonable to adopt these probabilities as intersubjective beliefs of the outcomes given the parameters (this is what makes the argument 3 in section 2.1 so important). The frequentist interpretation of the likelihood function implies that empirical tests with the data can be done. At the correct parameter values, the data should be a sample of the distribution formulated by the likelihood function. We cannot test this statement exactly, as we do not know the correct parameter values. Similarly as in frequentist statistics, we suggest to test the sample properties at the best guess for the correct parameter values, the maximum of the posterior (in frequentist statistics the best estimate is often given by the maximum of the likelihood function given the data). In many cases this will result in a residual analysis at the maximum of the posterior (see e.g. Yang et al. [2007]). Please note that we do not expect the observations to be a sample of the posterior. This is sometimes mistaken in the literature [e.g. Thyer et al. 2009]; due to the additional uncertainty resulting from uncertain knowledge of parameter values, the posterior should be systematically wider than the distribution of which the observations are a sample.

Residual analysis often leads to the insight, that a description of the residuals by a simple model of the observation error is incorrect. Residuals are very often larger and have a much stronger autocorrelation than expected for such an observation error model. This is an indication of model deficiencies that lead to systematic deviations or bias in model results that can hardly be eliminated completely. For this reason, it is very important to consider this bias explicitly in environmental modeling. The simplest way to do this is by using the statistical bias description technique due to Kennedy and O'Hagan [2001] (see also Bayarri et al. [2007] for a comprehensive description and Reichert and Schuwirth [2012], Dietzel and Reichert [2012] for its application in environmental modelling). Besides applying this technique to environmental modeling, Reichert and Schuwirth [2012] link it also conceptually to multi-objective model calibration [Gupta et al., 1998; Yapo et al. 1998; Vrugt et al. 2003].

Parallel to using a statistical description of bias, one should try to reduce the bias by improving the model structure. The use of time-dependent parameters [Tomassini et al. 2009; Reichert and Mieleitner, 2009] can be a constructive way to identify potential reasons of model structure deficits. Such techniques can be combined with expert elicitation about the structure and function of the investigated system and their relation to the model. Finally, all relevant error sources should be addressed with an approach such as BATEA [e.g. Renard et al. 2011], but the description of the "remnant errors" should be done using an approach such as the "bias description" approach discussed above [Reichert and Schuwirth, 2012].

3.3 Prediction

Prediction uncertainty bounds should clearly distinguish between the representation of our knowledge about the true system response, about potential new observations, and, if a frequentist likelihood function was used, conditional predictions at the best parameter estimates. Only the latter can be used to check the validity of statistical assumptions of a frequentist likelihood function by residual analysis.

4 NUMERICAL IMPLEMENTATION

In this section, important aspects of the numerical implementation of Bayesian inference for environmental modelling is briefly reviewed.

Numerical Optimization. If the likelihood function can be evaluated relatively cheaply (maybe after having performed a potentially demanding simulation of a deterministic sub-model), approximating the posterior numerically should always start with applying a numerical optimization technique. Due to the often complicated structure of the posterior, ideally this should be a global optimization algorithm (see e.g. Duan et al. [1993], Trelea [2003], Nocedal and Wright [2006]).

Markov Chain Monte Carlo. If the likelihood function can be evaluated relatively efficiently, Markov Chain Monte Carlo (MCMC) techniques are usually best suited to get a sample from the posterior that can be used for further inference [Gelman et al. 2004; Gamerman, 2006]. The use of an adaptive sampling scheme may be advantageous to increase the performance (see e.g. Haario [2001], Atchade and Rosenthal [2005], Haario et al. [2006]).

Approximate Bayes Computation. Consideration of stochasticity in the model as recommended in section 3.1 will often lead to difficulties in evaluating the likelihood function of the model, while still allowing to sample from it. In this case, Approximate Bayesian Computation (ABC) techniques will gain in importance to get a sample from an approximate posterior [Marjoram et al. 2003; Beaumont, 2009; Toni et al. 2009].

Use of Emulators. If the computer code implementing the model is slow, emulation of its output may be an option to increase the numerical efficiency of Bayesian inference, sensitivity analysis or prediction uncertainty estimation [Kennedy and O'Hagen, 2001; O'Hagan, 2006]. In the recent years, the approaches to emulation have been extended to emulating dynamic models [Bhattacharya, 2007; Liu and West, 2009; Conti et al., 2009; Conti and O'Hagan, 2010; Young and Ratto, 2011; Reichert et al., 2011; Castelletti et al. 2012; Reichert and Albert, 2012.].

5 CONCLUSIONS

This paper gave an overview of conceptual and practical aspects of uncertainty in environmental modelling. I hope that it will stimulate a constructive discussion among scientists about this issue that will finally lead to the development of improved techniques for estimating, propagating and communicating uncertainty. There are still major research needs in adequately formulating the likelihood function of the model (in particular regarding addressing all sources of uncertainty, the consideration of stochasticity and bias, and finding the adequate level of complexity), and in improving the efficiency of numerical algorithms e.g. by developing adaptive MCMC techniques and emulators. With the postulated move to stochastic models, new algorithms, such as ABC, will considerably gain in importance.

ACKNOWLEDGMENTS

I profited from a huge number of stimulating discussion over the past 20 years that are gratefully acknowledged. In addition, I would like to thank Carlo Albert, Anne Dietzel, Dario del Giudice, David Machac, Andreas Scheidegger, Nele Schuwirth and Colombe Siegenthaler – Le Drian for their comments to the manuscript.

REFERENCES

Abbaspour, K.C., Johnson, C.A. and van Genuchten, M. Th., Estimating uncertainty flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone Journal* 3:1340–1352, 2004.

Atchade, Y. F. and Rosenthal, J. On adaptive Markov chain Monte Carlo algorithms. *Bernoulli*, 11(5):815-828, 2005.

Bayarri, M. J., Berger, J. O., Paulo, R., Sacks, J., Cafeo, J. A., Cavendish, J., Lin, C.-H. and Tu, J. A framework for validation of computer models. *Technometrics*, 49(2), 138-154, 2007.

Beaumont M.A., Adaptive approximate Bayesian computation, *Biometrika* 96, 4, 983-990, 2009.

Beven, K. and Freer, J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* 249, 11-29, 2001.

Bhattacharya, S., A simulation approach to Bayesian emulation of complex dynamic computer models. *Bayesian Analysis* 2, 783–816, 2007.

Castelletti, P., Galelli, S., Restelli, M. and Soncini-Sessa, R., Data-driven dynamic emulation modelling for the optimal management of environmental systems, *Environmental Modelling & Software*, in press.

Clark, M.P. and Kavetski, D. Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of time stepping schemes. *Water Resources Research* 46, W10510, doi:10.1029/2009WR008894, 2010.

Colyvan, M., The philosophical significance of Cox's theorem. *Int. J. Approx. Reason*. 37, 71-85, 2004.

Colyvan, M. Is probability the only coherent approach to uncertainty? *Risk Analysis* 28(3), 645-652, 2008.

Conti, S., Gosling, J.P., Oakley, J., O'Hagan, A., Gaussian process emulation of dynamic computer codes. *Biometrika* 96, 663–676, 2009.

Conti, S. and O'Hagan, A., Bayesian emulation of complex multi-output and dynamic computer models. *Journal of Statistical Planning and Inference* 140, 640-651, 2010.

Cox, R.T. Probability, Frequency and Reasonable Expectation. *American Journal of Physics* 14(1), 1-13, 1946.

Dempster, A.P. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.* 38, 325-339, 1967.

DeRobertis, L. and Hartigan, J.A. Bayesian inference using intervals of measures. *The Annals of Statistics* 9(2): 235-244, 1981.

Dietzel, A. and Reichert, P. Calibration of computationally demanding and structurally uncertain models with an application to a lake water quality model. *Environmental Modelling & Software*, in review, 2012.

Duan, Q. Y., Gupta, V.K., and Sorooshian, S. Shuffled complex evolution approach for effective and effient global minimization. *Journal of Optimization Theory and Applications* 76 (3), 501-521, 1993.

Dubois, D. Possibility theory and statistical reasoning. *Computational Statistics & Data Analysis* 51, 47 – 69, 2006.

Dubois, D. and Prade, H. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988.

Gamerman, D. *Markov Chain Monte Carlo - Statistical Simulation for Bayesian Inference* (Second ed.). London: Chapman & Hall, 2006.

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B. *Bayesian Data Analysis* (Second ed.). London: Chapman & Hall, 2004.

Gillies, D. Intersubjective Probability and Confirmation Theory. *Brit. J. Phil. Sci.* 42: 513-533, 1991.

Gillies, D. *Philosophical Theories of Probability*. Routledge, London, UK, 2000.

Gupta, H. V., Sorooshian, S., and P. O. Yapo. P.O.. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research* 34 (4), 751-763, 1998.

Haario, H., Saksman, E., and Tamminen, J. An adaptive Metropolis algorithm. *Bernoulli*, 7(2):223-242, 2001.

Haario, H., Laine, M., Mira, A., and Saksman, E. DRAM: Effient adaptive MCMC. *Statistics and Computing*, 16:339-354, 2006.

Helton, J.C. and Oberkampf, W.L. (2004). Alternative representations of epistemic uncertainty. *Reliability Engineering and System Safety* 85, 1-10.

Howson, C. and Urbach, P. *Scientific Reasoning: The Bayesian Approach*. Open Court, La Salle, Illinois, USA, 1989.

Kavetski, D. and Kuczera, G. Model smoothing strategies to remove microscale discontinuities and spurious secondary optima in objective functions in hydrological calibration. *Wat. Resourc. Res.* 43, W03411, doi:10.1029/2006WR005195, 2007.

Kavetski, D. and Clark, M.P. Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction. *Water Resources Research* 46, W10511, doi:10.1029/2009WR008896, 2010

Kennedy, M. C. and O'Hagan, A. Bayesian calibration of computer models. *J. R. Statist. Soc. B: Statistical Methodology*, 63(3):425-464, 2001.

Lin, Z. and Beck, M.B. Accounting for structural error and uncertainty in a model: an approach based on model parameters as stochastic processes. *Environmental Modelling and Software* 27-28, 97-111, 2012.

Liu, F., West, M., A dynamic modelling strategy for Bayesian computer model emulation. *Bayesian Analysis* 4, 393–412, 2009.

Marjoram, P., Molitor, J., Plagnol, V. and Tavara, S. Markov chain Monte Carlo without likelihoods. *PNAS* 100(26), 15324-15328, 2003.

Moore, R.E. (1979). *Methods and Applications of Interval Analysis*. SIAM. Philadelphia.

Nocedal, J. and S. J. Wright. *Numerical Optimization* (second ed.). Springer, 2006.

Oberkampf, W.L., Helton, J.C., Joslyn, C.A., Wojtkiewicz, S.F. and Ferson, S. Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering and System Safety* 85, 11-19, 2004.

O'Hagan, A. Bayesian analysis of computer code outputs: a tutorial. *Reliability Engineering & System Safety* 91, 1290–1300, 2006.

Reichert, P. and Borsuk, M.E., Does high forecast uncertainty preclude effective decision support?, *Environmental Modelling and Software* 20(8), 991-1001, 2005.

Reichert, P., Langhans, S., Lienert, J. and Schuwirth, N., The Theoretical Foundation of Environmental Decision Support and its Application to River Management, in preparation.

Reichert, P., and Mieleitner, J., Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters, *Water Resources Research* 45, W10402, doi:10.1029/2009WR007814, 2009.

Reichert, P. and Schuwirth, N. Linking statistical description of bias to multi-objective model calibration. *Water Resources Research*, in review, 2012.

Reichert, P., White, G., Bayarri, M.J. and Pitman, E.B., Mechanism-based emulation of dynamic simulation models: Concept and application In hydrology. *Computational Statistics and Data Analysis*, 55, 1638–1655, 2011.

Reichert, P. and Albert, C. On the emulation of dynamic models. In preparation.

Renard, B., Kavetski, D., Leblois, E., Thyer, M., Kuczera, G. and Franks, S.W. Towards a reliable decomposition of predictive uncertainty in hydrological modelling: Characterizing rainfall errors using conditional simulation. *Water Resources Research* 47, W11516, doi:10.1029/2011WR010643, 2011.

Rinderknecht, S. L., Borsuk, M. E. and Reichert, P. Bridging Uncertain and Ambiguous Knowledge with Imprecise Probabilities, *Environ. Modelling & Software*, in press, 2011.

Rinderknecht, S.L., Borsuk, M.E. and Reichert, P. Eliciting Density Ratio Classes. *International Journal of Approximate Reasoning* 52, 792-804, 2011.

Rinderknecht, S.L., Albert, C., Schuwirth, N., Borsuk, M.E., Künsch H.-R. and Reichert, P., The Effect of Imprecise Prior Knowledge on Parameter Estimates and Model Predictions, in preparation, 2012.

Shafer, G. *A Mathematical Theory of Evidence*. Princeton Univ. Press. Princeton, 1976.

Snow, P. On the correctness and reasonableness of Cox's Theorem for finite domains. *Computational Intelligence* 14(3), 452-459.

Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S.W. and Srikanthan, S. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. *Water Resources Research*, 45, W00B14, doi:10.1029/2008WR006825, 2009.

Tomassini, L., Reichert, P., Künsch, H.-R. Buser, C., Knutti, R. and Borsuk, M.E., A smoothing algorithm for estimating stochastic, continuous-time model parameters and its application to a simple climate model, *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 58, 679-704, 2009.

Toni T., Welch D., Strelkowa N., Ipsen A., Stumpf M. P. H., Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems, *J. Roy. Soc. Int.* 6, 187-202, 2009.

Trelea, I. C. The particle swarm optimization algorithm: convergence analysis and parameter selection. *Information Processing Letters* 85, 317-325, 2003.

Vrugt, J.A., Gupta, H.V., Bastidas, L.A., Bouten, W. and Sorooshian, S. (2003). Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resources Research* 39 (8), Article Number 1214, doi:10.1029/2002WR001746.

Vrugt, J.A., Diks, C.G.H., Gupta, H.V., Bouten, W. and Verstraten, J.M. Improved treatment of uncertainty in hydrologic modelling: Combining the strength of global optimization and data assimilation. *Water Resources Research* 41, W01017, doi:10.1029/2004WR003059, 2005.

Walley, P. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, 1991.

Wilkinson, D.J. *Stochastic Modelling for Systems Biology*. 2nd ed. CRC Press, 2011.

Yang, J., Reichert, P., Abbaspour, K.C. and Yang, H., Hydrological modelling of the Chaohe Basin in China: Statistical model formulation and Bayesian inference, *Journal of Hydrology* 340, 167-182, 2007.

Yapo, P. O., Gupta, H.V. and Sorooshian, S. Multi-objective global optimization for hydrologic models. *Journal of Hydrology* 204, 83-97, 1998.

Young, P. and Ratto, M. Statistical emulation of large linear dynamic models. *Water Resources Research* 53(1), DOI 10.1198/TECH.2010.07151.

Zadeh, L.A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* 1, 3-28, 1978.